The generalizability of the structure of substance abuse and antisocial behavioral syndromes

Soe-Agnie, S E; Paap, M C S; VanDerNagel, J E L; Nijman, H. J. M.; de Jong, C. A. J.

Published in:
Psychiatry Research

DOI:
10.1016/j.psychres.2017.11.007

Publication date:
2018

Psychiatry Research 259 (2018) 412–421

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

The generalizability of the structure of substance abuse and antisocial behavioral syndromes: A systematic review

S.E. Soe-Agnie[a,*], M.C.S. Paap[b,c], J.E.L. VanDerNagel[d], H.J.M. Nijman[e], C.A.J. de Jong[d]

[a] Behavioural Science Institute BSI, Faculty of Clinical Psychology, Radboud University, The Netherlands
[b] Centre for Educational Measurement at the University of Oslo (CEMO), Faculty of Educational Sciences, University of Oslo, Oslo, Norway
[c] Department of Special Needs, Education, and Youth Care, Faculty of Behavioural and Social Sciences, University of Groningen, The Netherlands
[d] Nijmegen Institute for Scientist-Practitioners in Addiction (NISPA), Behavioral Science Institute, Faculty of Clinical Psychology, Radboud University, Nijmegen, The Netherlands
[e] Aventurijn Department at Altrecht GGZ, Den Dolder, The Netherlands

ARTICLE INFO

Keywords:
Factor structure
Antisocial
Substance abuse
Systematic review

ABSTRACT

Background: Although several authors have suggested that a single externalizing spectrum encompassing both antisocial behavioral syndromes and substance use disorder is to be preferred, this assumption has not been evaluated systematically throughout studies.
Purpose: The objective was to establish the generalizability of transdiagnostic models of externalizing disorders across different types of disorders and populations, in regard to the strength of the evidence.
Method: We conducted a systematic literature review using combinations of two sets of keywords: 1) "antisocial", "externalizing", "conduct disorder", "disruptive behavior disorder", "substance abuse", "substance-related disorder", "cannabis", "cocaine", "hallucinogen", "alcoholism", "opioid"; 2) "latent structure", "factor analysis", "multivariate analysis".
Results: Models supporting a superordinate factor appeared dominant in a limited set of different populations, on which the majority of the research sample was focused.
Conclusions: Although the externalizing spectrum model is a promising angle for future research and treatment, extending research on this model in a higher diversity of populations is recommended to enhance the understanding and applicability of the externalizing spectrum model.

1. Introduction

Substance use disorders (SUDs) and antisocial behavioral syndromes (ABSs; e.g., antisocial personality disorder and conduct disorder) are frequently co-occurring mental disorders, which are both associated with treatment drop-out, relapse in substance abuse and criminal recidivism (Daughters et al., 2008; Dykstra et al., 2015; Goldstein et al., 2007). A better understanding of the nature and relation of co-occurring disorders would benefit treatment in clinical practice. For instance, identifying overlapping core features and examining the complex interconnections among co-occurring disorders could make a more nuanced formulation of the diagnoses possible. This would lead to a more integrated treatment of these disorders. In order to achieve this, an accurate conceptualization of core features is essential. The externalizing spectrum model (Krueger, 1999) is a conceptual model that integrates alcohol dependence, drug dependence, and antisocial personality into a single model, taking into account the correlations among the different disorders. The term *externalizing disorders* was introduced by Achenbach and Edelbrock (1978), who used it to describe different forms of antisocial behaviors in children, not including substance abuse. The externalizing spectrum (Krueger, 1999) has been replicated in multiple studies (e.g., Slade and Watson, 2006; Vollebergh et al., 2001) and has been proposed for DSM-5 to represent one cluster in a higher-order meta-structure (Carragher et al., 2015; Krueger et al., 2005; Krueger and South, 2009). Although this model has not been included in its entirety in the final version of the DSM-5 (APA, 2014), several aspects have been adopted: Substance use disorders are now presented as a single structure, and antisocial personality disorder is now mentioned both in the Personality Disorder section and in the Disruptive Behaviors Disorders section of DSM-5. Furthermore, antisocial personality disorder (APD) is described as an externalizing disorder along with oppositional defiant disorder (ODD), conduct disorder

* Correspondence to: Nijmegen In stitute for Scientist-Practitioners in Addiction, Behavioral Science Institute, Radboud University, Toernooiveld 5, 6525 ED Nijmegen, The Netherlands.
E-mail address: sabrina.soeagnie@gmail.com (S.E. Soe-Agnie).

https://doi.org/10.1016/j.psychres.2017.11.007
Received 11 May 2017; Received in revised form 29 September 2017; Accepted 3 November 2017
Available online 04 November 2017
0165-1781/

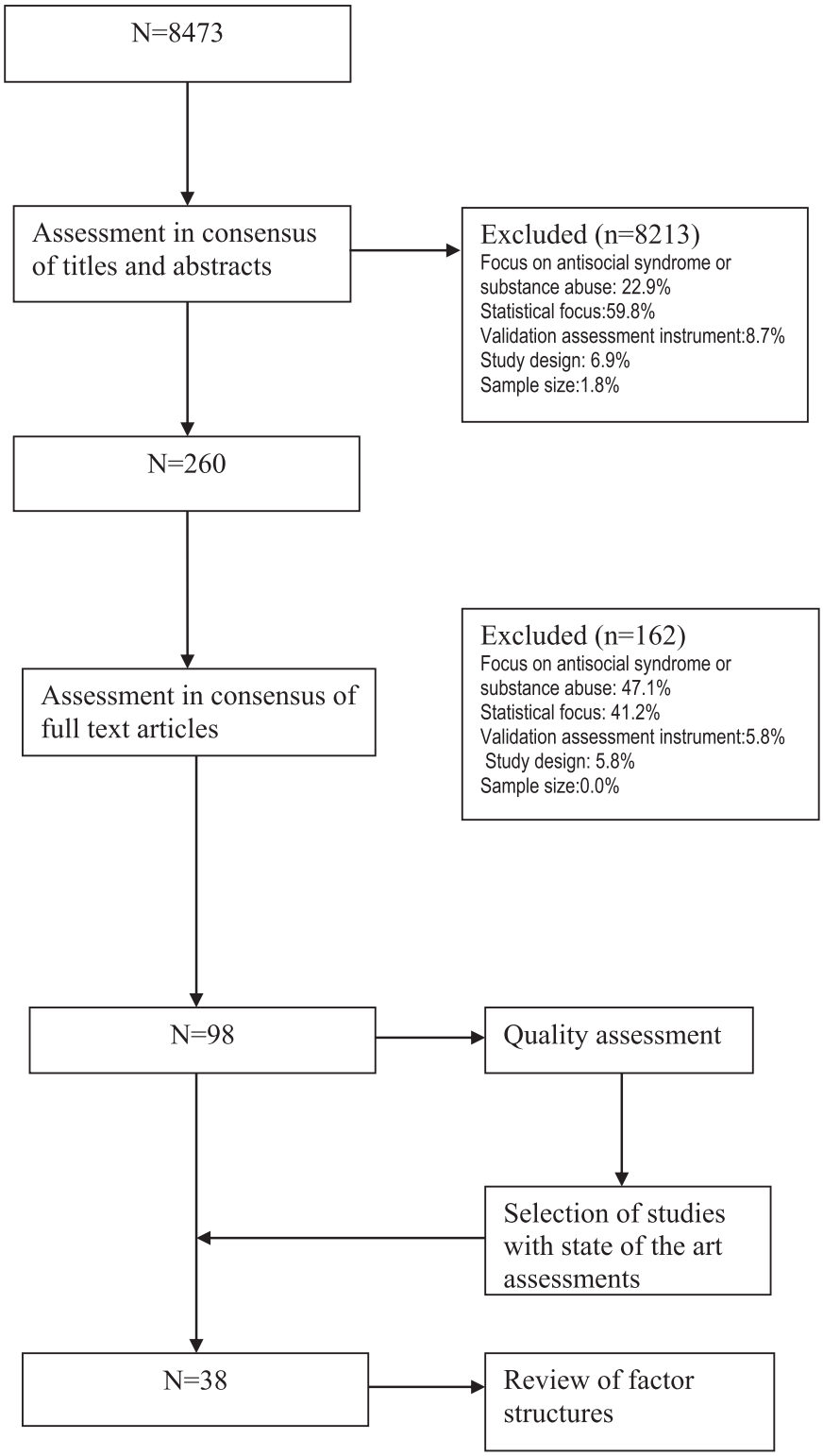

Fig. 1. Selection procedure.

(CD), and substance use disorders (SUDs) (APA, 2014). However, even though several factor analytic studies have found evidence supporting an externalizing spectrum model, the literature lacks a systematic quality assessment of these studies and is unclear on the generalizability of the externalizing model. Other reviews of the externalizing spectrum have solely focused on arguments supporting the concept without critical appraisal of the factor analytic studies supporting this spectrum model (e.g., Carragher et al., 2015; Eaton et al., 2015; Krueger and South, 2009). This is problematic, since features such as sample size, the level of heterogeneity of the sample and sampling bias could influence the interpretation of the applicability of a factor structure in different populations (Beavers et al., 2013; Gorsuch, 2015; MacCallum et al., 2001). Furthermore, it is unclear to what extent the generalizability of the externalizing spectrum is consistent across different samples and types of problem behaviors (e.g., different forms of substance use disorders or antisocial behaviors). For example, in scientific literature it is not clear to what extent Krueger's concept of externalizing disorders applies to child pathology, since research on youth

Table 1
Interrater reliability and consensus ratings of the quality assessment items (N = 98).

Quality items	κ (95%CI)	Response response categories (qual. points)	Consensus ratings %
Validity method			
Was the sampling method random?	.33 (.19–.46)	○ Random (1) ○ Quasi random (.5) ○ Convenience (0) ○ Unclear (0)	○ 22.2 ○ 15.2 ○ 24.3 ○ 38.4
Was an adequate sampling frame applied?	.28 (.15–.42)	○ Low bias (1) ○ Medium bias (.5) ○ High bias (0) ○ Unclear (0)	○ 25.3 ○ 14.1 ○ 24.2 ○ 36.4
Was the response rate adequate?	.65 (.53–.78)	○ > 70% (1) ○ 50–70% (.5) ○ < 50% (0) ○ Unclear (0)	○ 35.4 ○ 15.2 ○ 4.0 ○ 45.5
Measurement instruments			
a) Were the measurement instruments "state of the art"?	.68 (.56–.81)	○ Yes (1) ○ No, but common (.5) ○ No (0) ○ Unclear (0)	○ 36.4 ○ 9.153.5 ○ 1.0
a) Not "golden standard"/ common: Were the instruments reliable and valid?	.23 (− .05 to .50)	○ High (1) ○ Moderate (.5) ○ Low (0) ○ Unclear (0)	○ .0 ○ 7.3 ○ 1.8 ○ 90.9
Did the sample size meet the "rules of thumb" for factor analysis?	.31 (.04–.66)	○ At least 10 times number of items (1) ○ 5–10 times number of items (.5) ○ Less than 5 times number of items (0) ○ Unclear (0)	○ 90.9 ○ 5.1 ○ 2.0 ○ 2.0
Was a missing data analysis and (if relevant) a deletion/ imputation strategy described?	.26 (.11–.42)	○ Yes (1) ○ Partly (.5) ○ No (0) ○ Unclear (0)	○ 3.0 ○ 9.1 ○ 84.8 ○ 3.0
Was detailed information on study subjects provided?	.32 (.18–.47)	○ Yes (1) ○ Partly (.5) ○ No (0) ○ Unclear (0)	○ 51.5 ○ 31.3 ○ 12.1 ○ 5.1
Was detailed information on non-responders provided?	.27 (.13–.41)	○ Yes (1) ○ Partly (.5) ○ No (0) ○ Unclear (0)	○ 5.1 ○ 8.1 ○ 62.6 ○ 22.2

κ (95%CI) = Cohen's kappa interrater reliability with 95% confidence interval; consensus ratings = number of times a rating has been assigned after the consensus procedure; Response categories: ratings are ranked from low risk of bias to high risk of bias, followed by a rating of absence of information on the quality criterion. The quality points that correspond with each rating are presented between parenthesis.

externalizing behavior usually does not include substance abuse. Therefore, in order to establish the robustness and generalizability of the externalizing spectrum model across studies, a large-scale, systematic evaluation is required.

1.1. The present study

The first objective of this study is to systematically examine the strength of the evidence of factor models underlying externalizing disorders. The second objective is to explore if the transdiagnostic factor of externalizing disorders is generalizable across different samples and types of problem behaviors, including alcohol use disorder. So, we focused on studies describing different kinds of externalizing problem behaviors (different types of SUD, and ABS) and samples (clinical/community, adult/child, male/female) supporting a superordinate factor; here defined as either a *single factor*, *higher-order factor* or *strong general factor* underlying externalizing disorders.

2. Methods

2.1. Search strategy

We searched the literature using the search engines Pubmed, PsychInfo, and Embase, entering two sets of keywords and key phrases in quotation marks: 1) "antisocial", "externalizing", "conduct disorder", "disruptive behavior disorder", "substance abuse", "substance-related disorder", "cannabis", "cocaine", "hallucinogen", "alcoholism", "opioid"; 2) "latent structure", "factor analysis", "multivariate analysis".

The search was limited to peer-reviewed human studies, published in English between January 1990 and November 2014. This strategy initially resulted in a total of 8473 articles, which were evaluated by means of a structured selection procedure.

2.2. Selection procedure

Titles and abstracts of the initial selection of 8473 papers were studied and evaluated by two researchers, after which a consensus meeting was held to select articles for quality assessment, based on relevance and study design. After the consensus meeting, 260 studies were selected which met or roughly met the inclusion and exclusion criteria. The full-text copies of these studies were obtained, and the selection procedure was repeated. The criteria with regard to design characteristics (criteria 2–5) were chosen in order to retain factor analytic studies applying similar research designs. Articles were retained in the assessment procedure, if they met all of the following criteria:

1) The focus of the article was on antisocial personality, antisocial behavior, abuse or dependence of alcohol, legal or illegal drugs (the antisocial psychopathology was not linked to mental retardation or psychosis).
2) The statistical focus was either on principal component analysis, principal axis factoring, exploratory or confirmatory factor analysis.
3) The study was not focused on development and validation of new assessment instruments.
4) The study design was empirical, observational and quantitative, and the structural analyses were conducted cross-sectionally.
5) The sample size of the study had to exceed 100 subjects.

Based on the final evaluation, 98 studies were chosen for quality assessment. For further review of the factor structures, only the studies in which a state-of-the-art diagnostic assessment instrument was applied were included; e.g. (the Composite International Diagnostic Interview-Substance Abuse (CIDI-SAM; Cottler, 2000) or the Structured Clinical Interview for DSM-IV (SCID-II; First et al., 1997). This selection criterion was perceived as the most suitable for enabling a comparison among patterns of mental disorders. As a result, the factor structures of 38 articles were examined. The selection procedure is summarized in Fig. 1.

2.3. Quality assessment

The quality assessment was conducted between April 2011 and December 2012, and between November 2014 and August 2016. Since there was no standard method for assessing the quality of factor analytic studies, we designed this quality assessment instrument ourselves. The quality criteria in this study were based on criteria related to method validity and the generalizability of the results of epidemiologic studies by Loney et al. (1998), and The Strengthening the Reporting of Observational Studies in Epidemiology manual (Von Elm et al., 2007).

Table 2
Characteristics of included research samples and factor solutions categorized by SUD, ABS or SUDxABS.

Citation	Country	Gender	Age	Sampling	ABS type	SUD type	Objective Factor Analysis	Factors	No. qual.
SUD									
Beseler et al. (2006)	US	male	> 17	community	–	alcohol, cannabis, cocaine, hallucinogens, opiates, sedatives, stimulants	factor structure and biometrical modeling of symptoms of illicit substance abuse and dependence across seven substances	CFA: 1 factor for hallucinogens and 2 factors for remainders, with each factor containing different criteria	6.5
Borges et al. (2010)	Argentina Mexico, Poland, US	male		clinical	–	alcohol	factor structure of alcohol use disorders	EFA: 1 factor	4.5
Cherpitel et al. (2010)	Mexico, Poland, Argentina, US	male and female	> 17	clinical	–	alcohol	factor structure alcohol use disorder with and without craving criterion	EFA: 1 factor: F1 Dependence	3.0
[a] Feingold and Rounsaville (1995)	US	un-known	35.7	clinical and community	–	cocaine, alcohol, marijuana, opioids, sedatives, stimulants	factor structure of substance dependence syndrome across six substances	EFA: 1 factor: F1 Dependence syndrome	2.5
Gelhorn et al. (2008)	US	male and female	12–18	clinical and community	–	alcohol	factor structure alcohol abuse and dependence	EFA: 1 factor	2.0
Gillespie et al. (2007)	US	male and female	34	community	–	cannabis, cocaine, hallucinogens, sedatives, stimulants, opioids	factor structure 6 drug categories	CFA: 1 factor	3.0
Gillespie et al. (2011)	US	male and female	20–63	community	–	cannabis	factor structure cannabis abuse and dependence, comparison to LCA and FMM	CFA: 1 factor	2.5
Hartman et al. (2008)	US	male and female	15.2–16.3	community	–	cannabis	factor structure cannabis abuse and dependence	EFA: 1 factor	3.0
Kirisci et al. (2006)	US	male and female	–	community	–	alcohol, amphetamines, cannabis, cocaine, hallucinogens, inhalants, opioid, PCP, sedatives	factor structure SUD	CFA: 1 factor	1.5
Krueger et al. (2004)	US	male	43	community	–	alcohol	factor structure alcohol problems	PFA: 1 factor	6.0
Langenbucher et al. (2004)	US	male and female	35.3	clinical	–	alcohol, cannabis, cocaine	factor structure substance abuse and dependence across three substances	EFA: 1 factor for cocaine, 2 factors for rest: F1: All criteria except A3, D1; F2: A3, D1	3.0
Lewinsohn et al. (1996)	US	male and female	14–18	community	–	alcohol	factor structure of abuse and dependence	PCA&CFA: 2 factors: F1 Dependence withoutD6; F2 Abuse plus D6	4.5
Martin et al. (2006)	US	male and female	16.6	clinical	–	alcohol, cannabis	factor structure alcohol and cannabis use disorders	CFA: 1 factor	3.5
McBride et al. (2011)	Australia	male and female	45.8	community	–	alcohol	factor structure of alcohol use disorder and comparison to LCA and FMM	CFA: 1 factor	5.0
Mewton et al. (2011a)	Australia	male and female	> 17	community	–	alcohol	factor structure of alcohol use disorder	CFA: 1 factor	3.0
Mewton et al. (2011b)	Australia	male and female	> 17	community	–	alcohol	factor structure alcohol use disorder	CFA: 1 factor	4.5
Morgenstern et al. (1994)	US	male and female	36.1	clinical	–	alcohol, cannabis, cocaine, stimulants, opiates, sedatives, hallucinogens	factor structure of substance dependence syndrome across seven substances	PCA&CFA: 1 factor, except for hallucinogens	4.0
Nelson et al. (1999)	Netherlands, Turkey, Greece, Luxembourg, US, India	male and female		clinical and community	–	alcohol, cannabis, cocaine, opiates	factor structure substance disorder across four substances	EFA & CFA: 1 factor	2.0
Perron et al. (2010)	US	male and female	15.6	clinical	–	inhalants	factor structure inhalants use disorder	CFA: 1 factor	5.0
Piontek et al. (2011)	France	male and female	17–19	community	–	cannabis	factor structure cannabis abuse and dependence	CFA: 1 factor	5.5
Proudfoot et al. (2006)	Australia	male and female	> 17	community	–	alcohol	factor structure alcohol dependence	CFA: 1 factor	4.5

(*continued on next page*)

Table 2 (*continued*)

Citation	Country	Gender	Age	Sampling	ABS type	SUD type	Objective Factor Analysis	Factors	No. qual.
Rounsaville et al. (1993)	US	male and female	> 17	clinical and community		alcohol, cocaine, opiate, stimulant, sedative, cannabis	factor structure of substance dependence across six substances	EFA: 1 factor: F1 Dependence	4.0
Schafer and Caetano (1996)	US	male	> 17	clinical		cocaine	factor structure cocaine dependence	CFA: 1 factor: F1: dependence	4.0
Slade and Watson (2006)	Australia	male and female		community		alcohol, drugs	factor structure of 10 common DSM-IV and 11 common ICD-10 mental disorders including externalizing and internalizing disorders	EFA & CFA: 1 factor as part of higher order internalizing-externalizing factor	3.5
Srisunapanont et al. (2012)	Thailand	male and female	15–59	community	–	alcohol	factor structure Alcohol dependence	CFA: 1 factor: F1 Dependence	5.0
Teesson et al. (2002)	Australia	male and female	> 17	community		cannabis	factor structure cannabis dependence	CFA: 1 factor	5.0
Vollebergh et al. (2001)	Netherlands	male and female	18–64	community	–	alcohol, drugs	factor structure of nine common mental disorder including internalizing and externalizing disorders	CFA: 1 factor as part of higher order internalizing-externalizing factor	4.5
ABS									
Aebi et al. (2013)	Switserland	male and female	13.85	community	ODD	–	factor structure of ODD	CFA: 2 factors: F1 Irritable; F2 Headstrong/hurtful 3 factors: F1 Irritable; F2 Headstrong; F3 Hurtful	4.0
Bezdijan et al. (2011)	US	male	14,0	community	ODD and CD	–	multi-level hierarchical approach of ADHD, ODD, and CD	Hierarchical PCA: 3-level higher order component	2.5
Cosgrove et al. (2011)	US	male and female	14.84	community	ODD and CD	–	factor of internalizing and externalizing disorders	EFA: 1 as part of an internalizing-externalizing two-factors structure	3.0
Gelhorn et al. (2009)	US	male and female	14.7	community	CD	–	factor structure CD	EFA&CFA: 1 factor: F1: CD	3.0
Reitz et al. (2005)	Netherlands	male and female	13.36	community	other ABS, externalizing behavior	–	factor structure of internalizing and externalizing disorders	CFA: 1 factor as part of higher order internalizing-externalizing factor	3.0
SUD&ABS									
Krueger et al. (1998)	New Zealand	male and female	18–21	community	CD	cannabis, alcohol	structure of 10 common mental disorder including internalizing and externalizing disorders	CFA: 1 factor as part of higher order internalizing-externalizing factor	6.0
Krueger (1999)	US	male and female	15–54	community	AP	alcohol, drugs	factor structure 10 common mental disorders including internalizing and externalizing disorders	CFA: 1 factor as part of higher order internalizing-externalizing factor	3.0
Miller et al. (2008)	US	male	45	clinical	AP	drugs, alcohol	factor structure of comorbidity of 7 common mental disorders including internalizing and externalizing disorders	CFA: 1 factor as part of higher order internalizing-externalizing factor	3.0
Moss and Lynch (2001)	US	male and female	15.7	clinical and community	CD and ODD	alcohol	factor structure of ADHD, CD and ODD and alcohol use disorder separately	CFA: 1 factor: F1Alcohol and 2 factors: F1 CD: F2 ODD	4.5
Roysamb et al. (2011)	Norway	male and female	28.2	community	AP, CD	alcohol, hard drugs, cannabis	Joint factor structure of DSM-IV axis I and II disorders	EFA & CFA: 1 factor as part of 4-factor model	4.0
Wittchen et al. (2009)	Germany	male and female	28.8	community	AP, CD	alcohol, drugs	factor structure of 25 mental disorders	EFA & CFA: 1 factor as part of 6-factor model	5.0

No. qual. = number of quality criteria that are met. A criterion fully met counts for 1 point, a criterion partly met counts for .5 point.
ABS = antisocial behavioral syndrome; ADHD = attention defiant hyperactivity disorder; AP = antisocial personality disorder; CD = conduct disorder; CFA = confirmatory factor analysis; FMM = factor mixture modeling; IRT = item response modeling; LCA = latent class analysis; ODD = oppositional defiant disorder; PCA = principal component analysis; PCP =; SUD = substance abuse disorder.
[a] In this study dependence was assessed with a state-of the-art measure, while substance abuse was not. The substance abuse items are therefore disregarded in the result; US = United States; EFA = exploratory factor analysis.

Table 3
Frequency counts of sample characteristics for superordinate and multiple factor solutions.

	Superordinate factor N ($n_{quality\ 3>}$)	Multiple factors N ($n_{quality\ 3>}$)
Country		
US	19 (15)	5 (5)
Other	16 (15)	1 (1)
Gender		
Male	6 (6)	1 (1)
Male & Female	28 (23)	5 (5)
Age group		
12–17	11 (9)	3 (3)
18 >	22 (20)	3 (3)
Sampling		
Clinical	13 (10)	3 (3)
Community	27 (22)	4 (3)
ABS	13 (11)	3 (3)
ODD	2 (1)	1 (1)
CD	7 (6)	0 (0)
AP	4 (4)	0 (0)
Other ABS	2 (2)	1 (1)
SUD	73 (53)	9 (9)
Alcohol	23 (19)	3 (3)
Cannabis	11 (9)	3 (2)
Cocaine	7 (4)	1 (1)
Opioids	6 (3)	1 (1)
Amphetamines	1 (0)	0 (0)
Sedatives	3 (1)	1 (1)
Stimulants	4 (3)	1 (1)
Hallucinogens	2 (1)	1 (1)
Hard drugs	1 (1)	0 (0)
Drugs	5 (1)	0 (0)
Factor models		
First order	26 (22)	7 (6)
Part of broader first order multiple factor model	2 (2)	–
Higher order	1 (0)	–
Part of broader higher order model	6 (6)	–
Factor analyses		
CFA	18 (17)	4 (3)
EFA	8 (7)	1 (1)
PFA	1 (1)	0 (0)
PCA & EFA	1 (1)	1 (1)
PCA & CFA	0 (0)	1 (1)
EFA &CFA	3 (1)	0 (0)
Hierarchical PCA	1 (0)	0 (0)
Single & Hierarchical CFA	1 (1)	0 (0)
Number of quality points		
1.0–2.5	5	1
3.0–5.0	26	5
5.5–8.0	4	1

$n_{quality\ 3>}$ = frequency when references with quality below 3 points were discarded.

After consulting a statistician, two quality items referring to the quality of the statistical factor analyses were added based on Tabachnick and Fidell (2005, pp. 607, 615). In Table 1, the quality items and response categories are listed. Following recommendations of the Cochrane Handbook for Systematic Reviews of Intervention, Version 5.1.0 of The Cochrane collaboration (2011), the response categories were qualitative, corresponding respectively with "high risk of bias", "medium risk of bias", and "low risk of bias", and "unclear". The latter category was applied, when no information on an item could be found in the article that was being assessed. The rating of each article was carried out in pairs by a total of four raters: One epidemiologist (JvdN), two psychologists (MP and SS), and one forensic health care specialist (HN). Prior to the ratings, each pair of raters participated in a training session, in which five references were assessed and a consensus meeting was held, in order to reach uniformity in the interpretation of the quality items. After the assessment procedure, the interrater reliability of the quality items (Table 1) was poor to moderate (Hallgren, 2012). Therefore, SS and JvdN discussed all ratings to reach consensus and to achieve a better understanding of the reasons for the low interrater reliability, following Hallgren (2012). Based on these meetings, the quality item "Was a target population described in detail?" was considered to be open to too many misinterpretations and was excluded from further review. In addition, two issues became clear. First, part of the ratings showed large disparities; for example, one rater appraised an item as a high risk of bias, while another appraised the same item as a low risk of bias. Second, a consistent pattern appeared, in which the level of expertise of the raters involved seemed to influence the ratings. With regard to the sampling items, for example, the epidemiologists' ratings were mostly chosen as final scores in the consensus meetings, while the psychologists' ratings were mostly chosen as final scores with regard to measurement items. When a large disparity existed among the scores, we achieved consensus in consensus meetings. The remaining ratings for the subsets of items were assigned by the most knowledgeable researcher among three sets (SS and JvdN; SS and MP; SS and HN). The scores of two quality items were determined by either SS or JvdN, the scores of two items were determined by SS, the score of one item by MP, and the score of one item was determined by either MP, JvdN or SS. A points system was used to assess the quality of the selected studies: One point was awarded to each item that was fully met ("low risk of bias"), half a point to each item that was partially met ("medium risk of bias"), and, zero points to "high risk of bias" or "unclear" items (Table 1). Since eight quality items were evaluated for each study, and one point was always assigned to the selection criterion (the item referring to the quality of the assessment instruments), the research quality score ranged from one to eight points. We considered the research quality of a study acceptable, if a minimum of 75% of the items were rated as "medium risk of bias", corresponding with a research quality score of 3.0 points. A high research quality corresponds with a minimum of 63% (five out of eight) of the quality items being rated as "medium risk of bias" in combination with a minimum of 37% of the quality items being rated as "low risk of bias", resulting in a minimum of 5.5 points.

2.4. Statistical analyses

We analyzed the interrater agreement of the quality items using κ (kappa) in SPSS 23. We interpreted κ conservatively: Above .80 was good, .79–.68 was moderate, and below .67 was poor (Hallgren, 2012).

3. Results

3.1. Quality assessment

A total of 62 SUD, 21 ABS and 15 SUDxABS publications were assessed. The consensus ratings are presented in Table 1. Overall, sample sizes – which play an important role in factor analysis – were large in most studies, potentially increasing the reliability of the results. However, the percentages of "unclear" scores were high in items regarding sample selection, sample description and description of the psychometric properties of the assessment instrument that were not considered golden standard. In addition, missing data strategies were rarely described. Regarding the 38 references that were finally selected for further review, the quality ratings were equivalent to the original sample of references. An overview of the quality of the evidence of these 38 references is presented in Table 3 and will be described in Section 3.3.2.

3.2. Description of included studies

The study characteristics are presented in Table 2. Research data were often derived from large-scale studies, such as the Australian National Survey of Mental Health and Wellbeing (e.g., Baillie and Teesson, 2010), and multiple twin studies such as the Minnesota Twin Family Study (e.g., Krueger et al., 2004). Factors were mostly extracted using confirmatory factor analysis procedures (CFA). Higher order

models were mostly extracted through CFA, and, in one case, through a hierarchical principal component analysis (HPCA; e.g., Goldberg, 2006). All applied factor analysis methods are listed in Table 3. The majority of the analyses were based on symptom-level data (28 references), whereas the remainder were based on data at the diagnostic level (10 references). The diagnostic level was mostly applied in the SUDxABS studies. Nine studies not only addressed externalizing disorders, but included other disorders (e.g., internalizing disorders) in their analyses as well. Regarding the SUD citations, 19 out of 27 studies included all 11 DSM-IV (APA, 2000) substance abuse and dependence criteria, while the other references focused on dependence symptoms only. The disorders incorporated ODD, CD, APD, ABS not specified, alcohol, unspecified drugs and cannabis use disorders. Regarding the nationality of the samples, the United States (US) were most prevalent among all studies, followed by Australia and different countries in Europe, Asia and South-America.

3.3. Factor structures

3.3.1. Description of factors

Thirty-five out of the 38 reviewed references supported superordinate factors of which four studies, in which multiple disorders were addressed, supported a superordinate factor for only a part of these disorders. As is presented in Tables 2, 3, studies that found support for a superordinate factor differed in the type of model that was used and whether or not all studies disorders were subsumed by the superordinate externalizing factor. Examples of models include unidimensional models (symptom level, factor loadings λ = .27–.99 with only one study reporting loadings below .40), one higher-order component model (symptom level, factor loadings on level-one λ = .36–.70; level-two λ = .31–.78; level-three λ = .33–.62), and an externalizing factor as part of a non-hierarchical multifactor of multiple DSM-IV classifications (diagnostic level, λ = .46–.86). An externalizing factor was also described as part of a higher order factor, containing two level-two factors, labeled “internalizing” and “externalizing”, and several level-one factors subsumed by the higher-order internalizing factor (diagnostic level, λ = .47–.92.; for the symptom level the factor loadings were not reported). The quality ratings ranged from unacceptable to good. In most SUD studies (89.9%), equal support was found for both one-factor SUD and the two-factor model for substance abuse/dependency, which showed comparable model fit. However, due to the high intercorrelations (.85–1.00) of the factors in the two-factor model, resulting from substantial overlap, the one-factor models were preferred as the most parsimonious solution (i.e., the one-factor model is preferred as the least complex of the two, since it is able to explain the model well with fewer parameters). Respecting models not supporting a superordinate of transdiagnostic externalizing factor, quality ratings ranged from poor to good. Multiple factors of externalizing disorders were found in four SUD studies, one ABS study, and one SUDxABS study.

3.3.2. Generalizability

In several of the included studies, the generalizability of factors was specifically addressed. However, the results were diverse and even contradictory in some cases. For instance, superordinate factors were found for alcohol use disorder, cocaine use disorder and ABS across different countries, ethnic groups in the US, and genders (Cherpitel et al., 2010; Reitz et al., 2005; Shafer and Caetano, 1996). The generalizability of a superordinate factor was found to only apply to alcohol, cannabis, cocaine, stimulants, and opiates but not to hallucinogens in one study; while in another study the generalizability of a superordinate factor was found to apply to hallucinogens but not to other substances (Beseler et al., 2006; Morgenstern et al., 1994). Lastly, the generalizability of a transdiagnostic SUDxABS factor was found to be limited to adult subjects (excluding children and adolescents) in Wittchen et al. (2009). In Table 3, the frequency of sample and design characteristics in studies yielding either a superordinate factor or multiple factors are aggregated and listed. When a study yielded a superordinate factor for part of the disorders, while yielding multiple factors for the remaining disorders, each disorder was categorized into the corresponding column. When the other sample and design characteristics in a study were similar for the subordinate-factor disorders and the multiple-factor disorders, they were categorized into both columns. For example, Beseler et al. (2006) categorized *hallucinogens* into the superordinate factor column, while the other substances were categorized into the multiple factor column. Since subjects suffering from hallucinogen use disorders and the other SUDs were all recruited from a male, adult community sample from the US, these sample features were categorized into both superordinate factor and multiple factors columns. The frequency counts as presented in Table 3 suggest that superordinate factors prevailed over multiple factors in samples from both the US and other countries; samples that consisted of both genders combined; samples of adults; samples of from community populations; and samples involving, in particular, alcohol use. Superordinate factors were prevalent among clinical samples, and samples involving cannabis use disorder and ABS. The differences in frequency between superordinate and multiple factors were less apparent in all other characteristics due to low frequencies for both types of models. When references with low quality points (1.0–2.5) were discarded, the frequencies diminished particularly for the superordinate factors, resulting in the disappearance of differences in frequency between superordinate and multiple factors for several specific substances and ABSs. From a methodological perspective, superordinate factors appeared more prevalent in studies applying CFA compared to other extraction procedures.

4. Discussion

In the current review, a heterogeneous sample of references was reviewed in order to explore the strength of the evidence supporting a superordinate factor underlying externalizing disorders and to evaluate the reproducibility of this superordinate factor across different types of populations and externalizing problem behaviors.

4.1. Strength of the evidence

We deemed the quality (defined as the method validity and the generalizability of the results and quality of data analysis) of the larger part of the reviewed studies to be acceptable to good, with regard to models supporting both externalizing factors and multiple factors. The high sample sizes that were found in most studies are consistent with well documented recommendations for factor extraction (e.g., Tabachnick and Fidell, 2005, p. 613). The accuracy of sampling and measurement influences the interpretability of factors and their generalizability (Gorsuch, 2015). We found that the large quantities of data were often gained at the expense of quality, resulting in, for instance, a relatively small amount of studies applying state-of-the-art assessment instruments.

4.2. Generalizability of transdiagnostic, superordinate factors and impact of methodological and statistical features of studies

Most of the reviewed studies only seemed to support the generalizability of transdiagnostic or superordinate factors of externalizing disorders across a subset of populations, and overall the results were inconsistent. This could partly be the result of the different factor analytic procedures that were applied, e.g., principal component analysis (PCA) or confirmatory factor analysis (CFA); since the choice of the particular method of factor extraction can profoundly influence the factor solution. This means that different factor extraction methods of the same construct can lead to different factor structures. For instance, PCA was found to overestimate factor loadings in comparison to factor

analysis in a simulation study, while principal factor and maximum likelihood analysis are assumed to produce similar factor solutions (Gorsuch, 2015; Kim, 2008; Widaman, 1993). Factor solutions produced by EFA are found to be poorly replicated by CFA analysis (Van Prooijen and Van Der Kloot, 2001). Since the chosen extraction method is of importance to the replicability and therefore to the generalizability of factor solutions, an examination of the robustness of superordinate factors throughout all the analyses used would have been desirable in the current study. However, this was not possible, due to the lack of overlap in the methods used in the reviewed studies; particularly, when only references with acceptable to high quality levels were considered. Therefore, the impact of factor extracting procedures with regard to the externalizing factor remains unclear. The impact of the chosen factor models (e.g., first order or higher-order models) remains unclear as well, because of the heterogeneity of the types of models, the heterogeneity of the level of measurement (diagnostic vs. symptom level), and the variety of disorders.

4.3. Future directions research

Regarding the quality of factor analytic studies, we developed a new quality assessment instrument for factor analytic studies in the current review, because such an instrument was not available, yet. Since the quality of data and analyses impacts factor solutions, their interpretability and generalizability, we believe that more research is required to extend this work. An important direction of such research would be the improvement of the interrater reliability by, for instance, extending training sessions until the interrater reliability has reached an acceptable level, as is recommended in Hallgren (2012). Furthermore, factor analysis is a popular means of developing a better understanding of the interrelationship among comorbid disorders. Since factor analysis can be applied in different ways (e.g., exploratory or confirmatory), encompasses many different types of models (hierarchical/bi-factor/higher-order models), can be applied using different estimation methods (e.g., weighted least squares, full-information maximum likelihood), and researchers do not always use the same type of observed variables (symptoms/diagnoses), investigating replicability of a specific factor structure through literature review has proven to be challenging. An important next step towards increasing insight in the generalizability of the externalizing spectrum model would be to standardize the procedure being used (in terms of assessment instruments and psychometric methods); or, if that is not achievable, at least create enough overlap in the methods being used in order to enhance comparability across studies. Particularly, examining and comparing the impact of applying first order versus higher order and bifactor models is an interesting direction of research, since these models have gained popularity in recent studies on the structure of mental disorders (e.g., Blanco et al., 2015; Kotov et al., 2017; Noordhof et al., 2015; Snyder et al., 2017). Additionally, we recommend a shift in focus of studies on the structure of externalizing disorders from alcohol use disorders, and the use of community samples and samples consisting of adults, to other SUDs, ABSs, the comorbidity between different SUDs and ABSs, and the use of samples consisting of clinical subjects and adolescents. This would enable researchers to conduct multi-group statistical analyses in order to examine the generalizability of an externalizing factor solution. Furthermore, we recommend more uniformity and synergy between youth and adult psychiatry. The current review shows that, although studies have been conducted in both youth and adult psychiatry, the conceptualization of the externalizing spectrum differs between these fields. In youth psychiatry, studies on externalizing behavior usually exclude substance abuse, while in adult psychiatry, ADHD, for instance, is commonly omitted in studies on externalizing disorders. Both youth and adult psychiatric studies on externalizing disorders would benefit from including both ADHD and substance abuse, as research suggests, they are strongly interrelated with externalizing disorders in both adult and adolescent populations (e.g., Clark et al., 2002; Rucklidge and Downs-Woolley, 2016). Lastly, we recommend that researchers should focus more on the generalizability of this structure as well as the appropriate factor models to investigate this spectrum, *before* investing in further development of the externalizing spectrum (along with other spectra) as a novel classification system; as is proposed in Krueger and South (2009) for the DSM-5, and is presented in, for instance, Kotov et al. (2017).

4.4. Clinical implications

The concept of an externalizing spectrum as an underlying structure of overlapping features of SUDs and ABSs seems to be a promising angle for understanding of the interaction among co-occurring disorders and symptoms (Krueger and South, 2009; Tackett, 2010). The results of the current study imply that conceptualizing SUDs and ABSs as part of an externalizing spectrum, as is described in the DSM-5, seems to be consistent across community populations consisting of adults (m/f) particularly suffering from alcohol use disorder. However, it remains unclear if conceptualizing SUDs and ABSs as part of an externalizing spectrum also holds for clinical populations, youth, and populations suffering from other externalizing disorders than alcohol use disorder. Since the externalizing spectrum is meant to be a transdiagnostic model expressing the association among different forms of SUD and ABS, and the mention of this spectrum in the DSM-5 implies the usefulness of this concept for clinical populations, this result is unsatisfactory from a clinical perspective. Particularly, more clarity regarding the generalizability in clinical populations is considered important, since these populations would substantially benefit from a clear conceptualization and subsequent improved understanding and treatment of externalizing disorders. A developmental perspective of externalizing disorders and treatment development, as was suggested by Tackett (2010), for instance, would also benefit greatly from a shift in focus towards synergy among conceptualizations of disorders among youth and adults.

4.5. Limitations

Several large-scale and well-published studies were not included in this review, due to the selection criterion of state-of-the-art assessment. In addition, the number of ABS and SUDxABS references was low, due to the preference for golden standard assessment. Although, as a result of our choice some interesting findings may have been overlooked, uniformity in conceptual operationalizations of externalizing disorders has been guaranteed. Particularly, in regard to the generalizability of a particular concept, this uniformity of conceptualization is essential for a meaningful comparison across such diverse publications as were selected for the present review.

5. Conclusions

In conclusion, considering the strength of the evidence, the transdiagnostic factor of externalizing disorders seems generalizable across multiple sample characteristics and alcohol use disorder. However, research on the factor structure of externalizing based on DSM-classifications has been mostly limited to a small subset of populations, greatly restricting the interpretation of generalizability. Extending this research to a higher diversity of populations and externalizing disorders is recommended to improve the understanding and applicability of the externalizing spectrum model.

References

Achenbach, T.M., Edelbrock, C.S., 1978. The classification of child psychopathology: a review and analysis of empirical efforts. Psychol. Bull. 85, 1275–1301.

Aebi, M., Plattner, B., Winler Metzke, C., Bessler, C., Steinhausen, H., 2013. Parent- and self-reported dimensions of oppositionality in youth: construct validity, concurrent validity, and the prediction of criminal outcomes in adulthood. J. Child Psychol. Psychiatry 54, 941–949.

American Psychiatric Association, 2000. Diagnostic and Statistical Manual of Mental Disorders (Text Revision). fourth ed. American Psychiatric Association, Washington, D.C.

American Psychiatric Association, 2014. Brief overview of the DSM-5 criteria. [Beknopt overzicht van de criteria (DSM-5).] Dutch translation of the Desk reference to the diagnostic criteria from DSM-5. Boom, Amsterdam, pp. 340, 455–456.

Baillie, A.J., Teesson, M., 2010. Continuous, categorical and mixture models of DSM-IV alcohol and cannabis use disorders in the Australian community. Addiction 105, 1246–1253.

Beavers, A.S., Lounsbury, J.W., Richards, J.K., Huck, S.W., Skolits, G.J., Esquivel, S.L., 2013. Practical considerations for using exploratory factor analysis in educational research. PARE 18, 1–13.

Beseler, C., Jacobson, K.C., Kremen, W.S., Lyons, M.J., Glatt, S.J., Faraone, S.V., Gillespie, N.A., Tsuang, M.T., 2006. Is there heterogeneity among syndromes of substance use disorder for illicit drugs? Addict. Behav. 31, 929–947.

Bezdijan, S., Krueger, R.F., Derringer, J., Malone, S., McGue, M., Iacono, W.G., 2011. The structure of DSM-IV ADHD, ODD, and CD criteria in adolescent boys: a hierarchical approach. Psychiatry Res. 188, 411–421.

Blanco, C., Wall, M.M., He, J.P., Krueger, R.F., Olfson, M., Jin, C.J., Burstein, M., Merikangas, K.R., 2015. The space of common psychiatric disorders in adolescents: comorbidity structure and individual liabilities. J. Am. Acad. Child Adolesc. Psychiatry 54, 45–52.

Borges, G., Ye, Y., Bond, J., Cherpitel, C.J., Cremonte, M., Moskalewicz, Swiatkiewicz, G., Rubio-Stipec, R., 2010. The dimensionality of alcohol use disorders and alcohol consumption in a cross-national perspective. Addiction 105, 240–254.

Carragher, N., Krueger, R.F., Eaton, N.R., Slade, T., 2015. Disorders without borders: current and future directions in the meta-structure of mental disorders. Soc. Psychiatry Psychiatr. Epidemiol. 50, 339–350.

Cherpitel, C.J., Guilherme, B., Ye, Y., Bond, J., Cremonte, M., Moskalewicz, J., Swiatkiewicz, G., 2010. Performance of a craving criterion in DSM alcohol use disorders. J. Stud. Alcohol Drugs 71, 674–684.

Clark, D.B., Vanyukov, M., Cornelius, J., 2002. Childhood antisocial behavior and adolescent alcohol use disorders. Alcohol Res. Health 26, 109–115.

Cosgrove, V.E., Rhee, S.H., Gelhorn, H.L., Boeldt, D., Corley, R.C., Ehringer, M.A., Young, S.E., Hewitt, J.K., 2011. Structure and etiology of co-occurring internalizing and externalizing disorders in adolescents. J. Abnorm. Child Psychol. 39, 109–123.

Cottler, L.B., 2000. Composite International Diagnostic Interview—Substance Abuse Module (SAM). Department of Psychiatry, Washington University School of Medicine, St. Louis, MO.

Daughters, S.B., Stipelman, B.A., Sargeant, M.N., Schuster, R., Bornovalova, M.A., Lejuez, C.W., 2008. The interactive effects of antisocial personality disorder and court-mandated status on substance abuse treatment dropout. J. Subst. Abus. Treat. 34, 157–164.

Dykstra, R., Schumacher, J.A., Mota, N., Coffey, S.F., 2015. Examining the role of antisocial personality disorder in intimate partner violence among substance use disorder treatment seekers with clinically significant trauma. Violence Women 21, 958–974.

Eaton, N.R., Rodriguez-Seijas, C., Carragher, N., Krueger, R.F., 2015. Transdiagnostic factors of pathology and substance use disorders: a review. Soc. Psychiatry Psychiatr. Epidemiol. 50, 171–182.

Feingold, A., Rounsaville, B., 1995. Construct validity of the dependence syndrome as measured by DSM-IV for different psychoactive substances. Addiction 90, 1661–1669.

First, M.B., Gibbon, M., Spitzer, R.L., Williams, J.B.W., Benjamin, L.S., 1997. Structured Clinical Interview for DSM-IV Axis II Personality Disorders, (SCID-II). American Psychiatric Press, Inc., Washington, D.C.

Gelhorn, H., Hartman, C., Sakai, J., Mikulich-Gilbertson, S., Stallings, M., Young, S., Rhee, S., Corley, R., Hewitt, J., Hopfer, C., Crowely, T., 2008. An item-response theory analysis of DSM-IV conduct disorder. J. Am. Acad. Child Adolesc. Psychiatry 47, 1329–1339.

Gelhorn, H., Hartman, C., Sakai, J., Mikulich-Gilbertson, S., Stallings, M., Young, S., Rhee, S., Corley, R., Hewitt, J., Hopfer, C., Crowely, T., 2009. An item-response theory analysis of DSM-IV conduct disorder. J. Am Acad. Child Adolesc. Psychiatry 38, 42–50.

Gillespie, N.A., Neale, M.C., Prescott, C.A., Aggen, S.H., Kendler, K.S., 2007. Factor and item-response analysis DSM-IV criteria for abuse of and dependence on cannabis, cocaine, hallucinogens, sedatives, stimulants and opioids. Addiction 102, 920–930.

Gillespie, N.A., Kendler, K.S., Neale, M.C., 2011. Psychometric modeling of cannabis initiation and use and the symptoms of cannabis use, dependence and withdrawal in a sample of male and female twins. Drug Alcohol Depend. 118, 166–172.

Goldberg, L.R., 2006. Doing it all bass-ackwards: the development of hierarchical factor structures from the top down. J. Res. Pers. 40, 347–358.

Goldstein, R.B., Compton, W.M., Pulay, A.J., Ruan, W.J., Pickering, R.P., Stinson, F.S., Grant, B.F., 2007. Antisocial behavioral syndromes and DSM-IV drug use disorders in the United States: results from the national epidemiologic survey on alcohol and related conditions. Drug Alcohol Depend. 90, 145–158.

Gorsuch, R.L., 2015. The Replication and Invariance of Factors. Factor Analysis: Classical Edition. Taylor & Francis, New York.

Hallgren, K.A., 2012. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor. Quant. Methods Psychol. 8, 23–34.

Hartman, C.A., Gelhorn, H., Crowley, T., Sakai, J.T., Stallings, M., Young, S.E., Rhee, S.H., Corley, R., Hewitt, J.K., Hopfer, C.J., 2008. Item response theory analysis of DSM-IV cannabis abuse and dependence criteria in adolescents. J. Am. Acad. Child Adolesc. Psychiatry 47, 165–173.

Kim, H., 2008. Common factor analysis versus principal component analysis: choice for symptom cluster research. Asian Nurs. Res. 2, 17–24.

Kirisci, L., Tarter, R.E., Vanyukov, M., Martin, C., Mezzich, A., Brown, S., 2006. Application of item response theory to quantify substance use disorder severity. Addict. Behav. 31, 1035–1049.

Kotov, R., Krueger, R.F., Watson, D., Achenbach, T.M., Althoff, R.R., Bagby, R.M., Brown, T.A., Carpenter, W.T., Caspi, A., Clark, L.A., Eaton, N.R., Forbes, M.K., Forbush, K.T., Golberg, D., Hasin, D., Hyman, S.E., Ivanova, M.Y., Lynam, D.R., Markon, K., Miller, J.D., Moffit, T.E., Morey, L.C., Mullins-Sweatt, S.N., Ormel, J., Patrick, C.J., Regier, D.A., Rescorla, L., Ruggero, C.J., Samuel, D.B., Sellbom, D.B., Simms, L.J., Skodol, A.E., Slade, T., South, S.C., Tackett, J.L., Waldman, I.D., Waszcuk, M.A., Widiger, T.A., Wright, A.G.C., Zimmerman, M., 2017. The hierarchical taxonomy of psychopathology (HiTOP): a dimensional alternative to traditional nosologies. J. Abnorm. Psychol. 126, 454–477.

Krueger, R.F., 1999. The structure of common mental disorders. Arch. Gen. Psychiatry 56, 921–926.

Krueger, R.F., Caspi, A., Moffitt, T.E., Silva, P.A., 1998. The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. J. Abnorm. Psychol. 107, 216–227.

Krueger, R.F., Markon, K.E., Patrick, C.J., Iacono, W.G., 2005. Externalizing psychopathology in adulthood: a dimensional-spectrum conceptualization and its implications for DSM-V. J. Abnorm. Psychol. 114, 537–550.

Krueger, R.F., Nichol, P.E., Hicks, B.M., Markon, K.E., Patrick, C.J., Iacono, W.G., McGue, M., 2004. Using latent trait modeling to conceptualize an alcohol problems continuum. Psychol. Assess. 16, 107–119.

Krueger, R.F., South, S.C., 2009. Externalizing disorders: cluster 5 of the proposed meta-structure for DSM-V and ICD-11. Psychol. Med. 39, 2061–2070.

Langenbucher, J.W., Labouvie, E., Martin, C.S., Sanjuan, P.M., Bavly, L., Kirisci, L., 2004. An application of item response theory analysis to alcohol cannabis, and cocaine criteria in DSM-IV. J. Abnorm. Psychol. 113, 72–80.

Lewinsohn, P.M., Rohde, P., Seeley, J.R., 1996. Alcohol consumption in high school adolescents: frequency of use and dimensional structure of associated problems. Addiction 91, 375–390.

Loney, P.L., Chambers, L.W., Bennet, K.J., Roberts, J.G., Stratford, P.W., 1998. Critical appraisal of the health research literature: prevalence or incidence of a health problem. Chronic Dis. Can. 19, 170–176.

MacCallum, R.C., Widaman, K.F., Preacher, K.J., Hong, S., 2001. Sample size in factor analysis: the role of model error. Multivar. Behav. Res. 36, 611–637.

Martin, C.S., Chung, T., Kirisci, L., Langenbucher, J.W., 2006. Item response theory analysis of diagnostic criteria for alcohol and cannabis use disorders in adolescents: implications for DSM-V. J. Abnorm. Psychol. 115, 807–814.

McBride, O., Teesson, M., Baillie, A., Slade, T., 2011. Assessing the dimensionality of lifetime DSM-iv alcohol use disorders and a quantity-frequency alcohol use criterion in the Australian population: a factor mixture modeling approach. Alcohol Alcohol. 46, 333–341.

Mewton, L., Slade, T., McBride, O., Grove, R., Teesson, M., 2011. An evaluation of the proposed DSM-5 alcohol use disorder criteria using Australian national data. Addiction 106, 941–950.

(b)Mewton, L., Teesson, M., Slade, T., Cottler, L., 2011b. Psychometric performance of DSM-IV alcohol use disorders in young adulthood: evidence from an Australian general population sample. J. Stud. Alcohol Drugs 72, 811–822.

Miller, M.W., Fogler, J.M., Wolf, E.J., Kaloupek, D.G., Keane, T.M., 2008. The internalizing and externalizing structure of psychiatric comorbidity in combat veterans. J. Trauma Stress 21, 58–65.

Moss, H.B., Lynch, K.G., 2001. Comorbid disruptive behavior disorder symptoms and their relationship to adolescent alcohol use disorders. Drug Alcohol Depend. 64, 75–83.

Morgenstern, J., Langenbucher, Labouvie, E.W., 1994. The generalizability of the dependence syndrome across substances: an examination of some properties of the proposed DSM-IV dependence criteria. Addiction 89, 1105–1113.

Nelson, C.B., Rehm, J., Üstün, T.B., Grant, B., Chatterji, S., 1999. Factor structures for DSM-IV substance disorder criteria endorsed by alcohol, cannabis, cocaine and opiate users: results from the WHO reliability and validity study. Addiction 94, 843–855.

Noordhof, A., Krueger, R.F., Ormel, J., Oldehinkel, A.J., Hartman, C.A., 2015. Integrating autism-related symptoms into the dimensional internalizing and externalizing model of psychopathology. The TRAILS study. J. Abnorm. Child Psychol. 43, 577–587.

Perron, B.E., Vaughn, M.G., Howard, M.O., Bohnert, A., Guerrero, E., 2010. Item response theory analysis of DSM-IV criteria for inhalant-use disorders in adolescents. J. Stud. Alcohol Drugs 71, 607–614.

Piontek, D., Kraus, L., Legleye, S., Bühringer, G., 2011. The validity of DSM-IV cannabis abuse and dependence criteria in adolescents and the value of additional cannabis use indicators. Addiction 106, 1137–1145.

Proudfoot, H., Baillie, A.J., Teesson, M., 2006. The structure of alcohol dependence in the community. Drug Alcohol Depend. 81, 21–26.

Reitz, E., Dekovic, M., Meijer, A.M., 2005. The structure and stability of externalizing and internalizing problem behavior during early adolescence. J. Youth Adolesc. 34, 577–588.

Rounsaville, B.J., Bryant, K., Babor, T., Kranzler, H., Kadden, R., 1993. Cross-system agreement for substance use disorders: DSM-III-R, DSM-IV and ICD-10. Addiction 88, 337–348.

Roysamb, E., Tambs, K., Orstavik, R.E., Torgersen, S., Kendler, K.S., Neale, M.C., Aggen, S.H., Reichborn-Kjennerud, T., 2011. The joint structure of DSM-IV axis I and axis II disorders. J. Abnorm. Psychol. 120, 198–209.

Rucklidge, J.L., Downs-Woolley, M., 2016. Psychiatric comorbidities in a New Zealand sample of adults with ADHD. J. Atten. Disord. 20, 1030–1038.

Shafer, J., Caetano, R., 1996. The DSM-IV construct of cocaine dependence in a treatment sample of black, Mexican American, and white men. Psychol. Assess. 8, 304–311.

Slade, T., Watson, D., 2006. The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. Psychol. Med. 36, 1593–1600.

Snyder, H.R., Young, J.F., Hankin, B.L., 2017. Chronic stress exposure, and generation are related to the P-factor and externalizing specific psychopathology in youth. J. Clin. Child Adolesc. Psychol. 00, 1–10.
Srisunapanont, M., Kittiratanapaiboon, P., Likhitsathian, S., Kongsuk, T., Suttajit, S., Junsirimongkol, B., 2012. Patterns of alcohol dependence in Thai drinkers: a differential item functioning analysis of gender and age bias. Addict. Behav. 37, 173–178.
Tackett, J.L., 2010. Toward an externalizing spectrum in DSM-V: incorporating developmental concerns. Child Dev. Perspect. 4, 161–167.
Tabachnick, B.G., Fidell, L.S., 2005. Using Multivariate Statistics. Pearson Education, Boston, pp. 5.
Teesson, M., Lynskey, M., Manor, B., Baillie, A., 2002. The structure of cannabis dependence in the community. Drug Alcohol Depend. 68, 255–262.
The Cochrane collaboration, 2011. Quality scales and cochrane reviews. In: Higgins, J.P.T., Green, S. (Eds.), Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0, (Accessed 1 April 2011). 〈http://www.handbook.cochrane.org〉.
Vollebergh, W.A.M., Iedema, J., Bijl, R.V., De Graaf, R., Smit, F., Ormel, J., 2001. The structure and stability of common mental disorders: the NEMESIS study. Arch. Gen. Psychiatry 58, 597–603.
Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gotzsche, P.C., Vandenbroucke, J.P., STROBE Initiative, 2007. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observation studies. J. Clin. Epidemiol. 61, 344–349.
Van Prooijen, J., Van Der Kloot, W.A., 2001. Confirmatory analysis of exploratory obtained factor structures. Educ. Psychol. Meas. 61, 777–792.
Widaman, K.F., 1993. Common factor analysis versus principal component analysis: differential bias in representing model parameters? Multivar. Behav. Res. 28, 263–311.
Wittchen, H.U., Beesdo-Baum, K., Gloster, A.T., Höfler, M., Klotsche, J., Lieb, R., Beauducel, A., Bühner, M., Kessler, R.C., 2009. The structure of mental disorder re-examined: is it developmentally stable and robust against additions? Int. J. Methods Psychiatr. Res. 18, 189–203.